

ATP binding site of P2X channel proteins: structural similarities with class II aminoacyl-tRNA synthetases

Wolfgang Freist, Janko F. Verhey, Walter Stühmer*, Dieter H. Gauss

Max-Planck-Institut für experimentelle Medizin, Hermann-Rein-Str. 3, D-37075 Göttingen, Germany

Received 26 July 1998

Abstract The extracellular loop of P2X channel proteins contains a sequence stretch (positions 170–330) that exhibits similarities with the catalytic domains of class II aminoacyl-tRNA synthetases as shown by secondary structure predictions and sequence alignments. The arrangement of several conserved cysteines (positions 110–170) shows similarities with metal binding regions of metallothioneins and zinc finger motifs. Thus, for the extracellular part of P2X channel proteins a metal binding domain and an antiparallel six-stranded β -pleated sheet containing the ATP binding site are very probable. The putative channel forming H5 part (positions 320–340) shows similarities with the enzyme motif 1 responsible for aggregation of subunits to the holoenzyme.

© 1998 Federation of European Biochemical Societies.

Key words: Ligand-gated channel; ATP receptor; Purinoceptor; P2X; Aminoacyl-tRNA synthetase

1. Introduction

ATP-gated ion channels (P2X) have been established as one of two main families of purinergic receptors. Today 13 of them have been cloned and sequenced (listed in [1–3]). The channels are permeable to monovalent and divalent cations, and possess two hydrophobic regions that are putative transmembrane domains. These are separated by an extracellular loop containing almost 300 amino acid residues. Sequence homology in this loop region among the various P2X channels is 50–60%. The three-dimensional structure of P2X channels is still unknown; neither X-ray nor NMR analyses have been published.

Aminoacyl-tRNA synthetases catalyse the esterification of a specific tRNA with a particular amino acid. The energy required for this process is obtained by conversion of ATP to AMP and PP_i. These enzymes can be divided into two groups of 10 enzymes each, defined as class I and class II [4–9]. The enzymes of the two groups are principally different by two distinct types of tertiary structures.

Class II synthetases exhibit three homologous sequences called motifs 1, 2 and 3. The latter two are regions of a specific tertiary fold built around a central sheet of six antiparallel β -sheet strands and a long α -helix [5]; they are spatially close and participate in forming the active centre that contains the ATP binding site (compare Fig. 1).

Enzymes specific for the same amino acid but isolated from different species show sequence homology in the range of 20–40% [10–12]. In contrast, homology among enzymes specific for different amino acids from different sources is poor. With-

in the partial sequences of motifs 1, 2, and 3 they can be estimated to approximately 10% (see [5]).

A series of crystal structures of class II aminoacyl-tRNA synthetases are currently known; these enzymes are specific for serine (*Escherichia coli* [10], *Thermus thermophilus* [13–17]), histidine (*E. coli* [18]), aspartic acid (*T. thermophilus* [19,20], yeast [21–23]), glycine (*T. thermophilus* [24]), phenylalanine (*T. thermophilus* [25–27]), and lysine (LysU from *E. coli* [28]).

In this article similarities between P2X ion channels and class II aminoacyl-tRNA synthetases are described, especially in those regions which may be responsible for ATP binding.

2. Materials and methods

Secondary structure predictions were made by the algorithms given in [29–33] for the individual channel proteins. In Fig. 1, together with the partial sequences displayed, the percentages of secondary structure predictions for β -sheet strands, helices or neither are given.

3. Results

In Fig. 1 aligned regions of P2X channel proteins, for which six β -sheet strands were predicted, are compared with those β -sheet strands of class II aminoacyl-tRNA synthetases found by X-ray analysis as regions of the ATP binding catalytic centre. According to the enzyme structures they are designated β_I – β_{VI} . In the enzymes β -sheet strands β_I and β_{II} belong to motif 2, whereas β_{VI} is a region of motif 3.

In the β_I region the channel proteins show a relatively well conserved sequence consisting of two hydrophobic residues, and additionally a lysine and an asparagine ($\Phi\Phi KN$) (Fig. 1). The aligned sequences of the enzymes have the common structure $\Phi\Phi RXE$. In addition to the hydrophobic residues, K and N of the channel sequences may correspond to R and E of the enzymes. All four columns are completely conserved and thus seem to be important. A consensus sequence may be formulated as $\Phi X_7 \Phi_2 (RK) X (EN)$. The arginines of the enzyme proteins participate in binding of the α -phosphate moiety of ATP. This function may also very well be fulfilled by the aligned lysine of the channel proteins. At a distance of about six residues towards the C-terminal, the channel proteins exhibit a column consisting of arginine, lysine, and histidine, whereas the aligned enzyme sequences show a column consisting of arginine and histidine. In the enzymes, these residues participate in binding of the γ -phosphate moiety of ATP, and this function may again be fulfilled by the corresponding residues of the channel proteins.

In the β_{II} region, a near constant FRLG sequence common to the channel proteins corresponds to similar sequences of the enzymes (for example FHQLG) (Fig. 1, β_{II}). These stretches are largely followed by hydrophobic and acidic

*Corresponding author. Fax: (49) (551) 3899644.
E-mail: stuhmer@mail.mpiem.gwdg.de

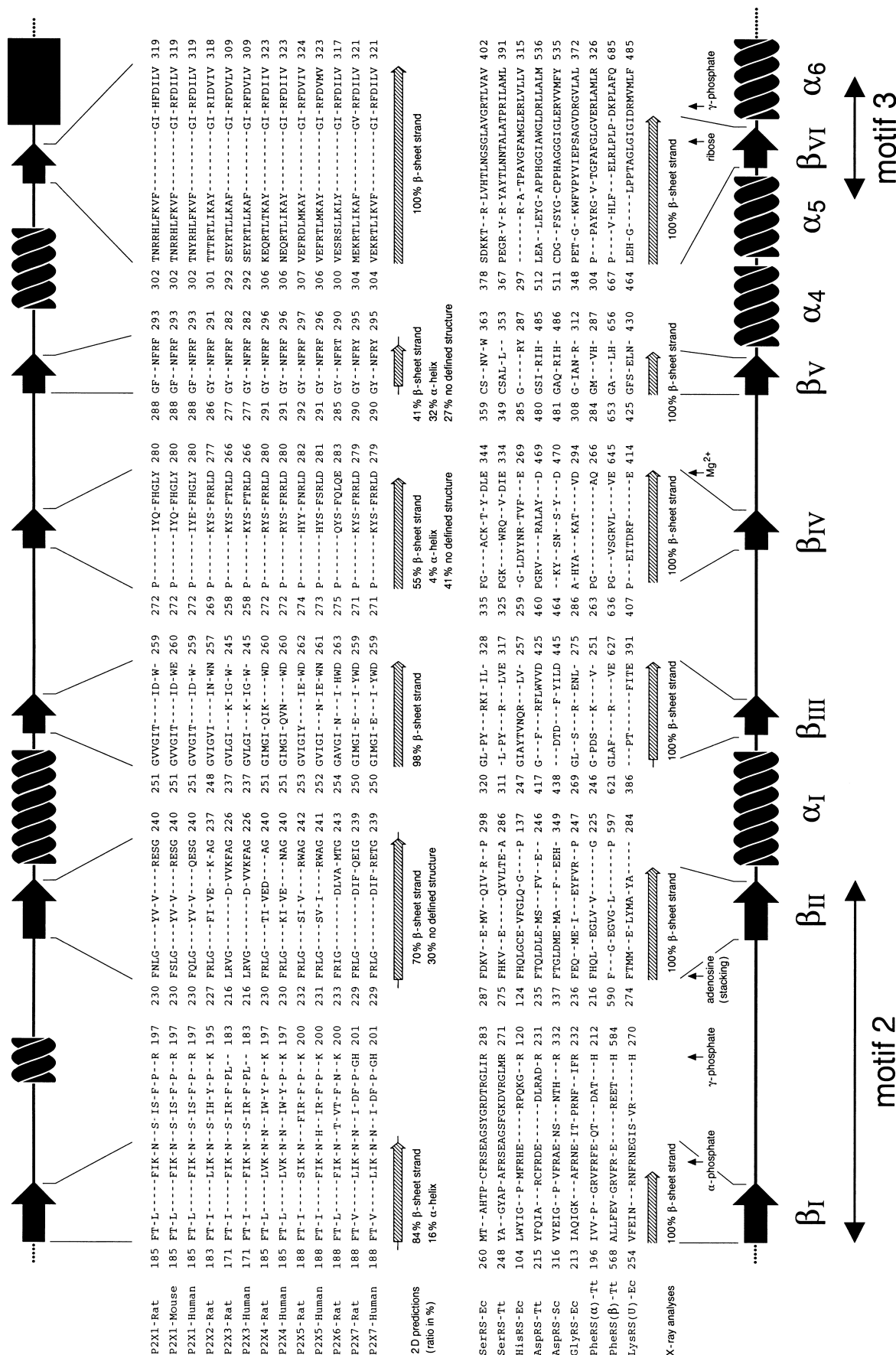


Fig. 1. Alignments of 13 characteristic sequence regions of P2X channel proteins (above) and nine class II aminoacyl-tRNA synthetases (below) which include partly or completely the β -sheet strand regions. The diagrams at the top and bottom indicate: arrow for β -sheet strand, helical structure for α -helix and box for transmembrane domains. Amino acid residues involved in binding of ATP by class II aminoacyl-tRNA synthetases are marked by arrows. The ratios of β -sheet strand or α -helix predictions obtained by five different prediction algorithms are given as percentage below the P2X channel sequences. β -Sheet regions of class II aminoacyl-tRNA synthetases are taken from crystal structures of the nine enzyme sequences. The Swiss-Prot and PIR database registration numbers are: P2X1-Rat P47824; P2X1-Mouse P51576; P2X1-Human P51575; P2X2-Rat P49653; P2X3-Rat P49654; P2X4-Rat P51577; P2X4-Human Q99571, O00450; P2X5-Rat P51578, Q64613; P2X5-Human Q93086, Q93087; P2X6-Rat P51579; P2X7-Rat Q64663; P2X7-Human Q64663; SerRS-Ec P09156, SerRS-Tt P34945, HisRS-Ec P04804, AspRS-Tt P36419, AspRS-Sc P04802, PheRS(α)-Tt P27001, PheRS(β)-Tt P27002, LysRS(U)-Ec P14825 (all Swiss-Prot) and GlyRS-Tt S58522 (PIR) (Ec = *Escherichia coli*, Tt = *Thermus thermophilus*, Sc = *Saccharomyces cerevisiae*).

sequences and a final G or P. Whereas the first region of this block shows convincing similarities, sequences are less similar in the second region. Nevertheless, a consensus sequence can be derived as follows:

F(RH)-X(Φ G) $X_4\Phi_2X\Phi$ (EDQ) Φ_3 (RK) X_2 (GP). In the enzymes, the conserved phenylalanine side chain is responsible for stacking interactions with the adenine of the ATP molecule, and a counterpart of this hydrophobic and aromatic residue can again be seen in P2X channel proteins.

In all tertiary structures of class II aminoacyl-tRNA synthetases a structure immanent α -helix part was found between β_{II} and β_{III} ; this structure is also predicted for all P2X channel proteins at the corresponding positions (not shown).

For the β_{III} region three characteristic parts may be mentioned (Fig. 1, β_{III}): in the channel proteins a conserved G Φ F Φ G stretch appears which is less well conserved in the enzymes. In the central part (NQ), (RK), and hydrophobic columns are present, though incomplete in both groups. Finally, at the end of this region columns of acidic residues in both protein groups are preceded by hydrophobic residues. In summary, although there is no residue conserved in all proteins, the general character of this region seems to be similar for both groups. This is also indicated by the suggested consensus sequence G Φ_2 (PG) X_2 (NQ)X(RK)X Φ X Φ_2 (DE).

In the β_{IV} region, the channel proteins again show more homology than the aminoacyl-tRNA synthetases, although they are not as well conserved as in the β_V and β_{VI} regions

(Fig. 1, β_{IV}). Commonly, both families start with P or (PG), and end with a D or a (DE) column. The central section comprises (KRH), (YST) and hydrophobic stretches, and the consensus sequence $PX_7(RKH)X(FY)X_2\Phi(ED)$ can be given.

The acidic residue at the end of the β_{IV} region has an important function in complexing of the magnesium ion of an [ATP-Mg] $^{2-}$ which is the actual substrate for the enzymes and probably also the agonist for P2X channels (Fig. 1, β_{IV}). A second acidic residue responsible for the same function and also for contacting the ribose moiety is situated between β -sheet regions β_{IV} and β_V of class II aminoacyl-tRNA synthetases. Again, corresponding acidic residues are found in the P2X sequences (not shown).

In the region of sheet strand β_V the aminoacyl-tRNA synthetases each possess a slightly differing sequence whereas the P2X sequences are nearly constant [GYNFRF] (Fig. 1, β_V). Common to both protein families are a G column in the first position, an N column and an R or (HR) column all followed by residues of predominantly hydrophobic character. For both protein families the consensus sequence G(SY) X_2 (NR) Φ (HR) Φ suggests a similarity.

The most evident analogy occurs in the C-terminal region of the β_{VI} -sheet strand and its counterpart of the channel proteins (Fig. 1, β_{VI}). In both protein families three typical features can be recognised: the first position of these stretches is largely occupied by G followed by a column of hydrophobic residues, in the central region conserved columns of R and D are clearly observed, and 3–5 aminoacyl residues with hydrophobic side chains occur at the end (Fig. 1, β_{VI}). Thus, a consensus sequence G Φ XRX Φ X Φ_3 is clearly indicated with the first X position occupied by acidic residues in the enzymes and the second X position by D in all channel sequences. For the N-terminal part of the β_{VI} region a consensus sequence X(DE)X(KR)X Φ X(KR) $\Phi\Phi$ can be postulated. Altogether, channel and enzyme proteins exhibit relatively conserved sequences within this β_{VI} strand region and both protein families correspond well.

Class II aminoacyl-tRNA synthetases use the conserved arginine of the C-terminal part of the β_{VI} region for the interaction with the γ -phosphate moiety of the ATP molecule. Hydrogen bonds are formed by the enzymes with the ribose part of ATP by a glycine of motif 3. These functions may be easily performed in the P2X channel proteins by the conserved arginines and glycines of the P2X channels in the C-terminal sequence.

4. Discussion

Aminoacyl-tRNA synthetases functioning as important parts of the machinery which translates the genetic information into protein structures must have been present in their

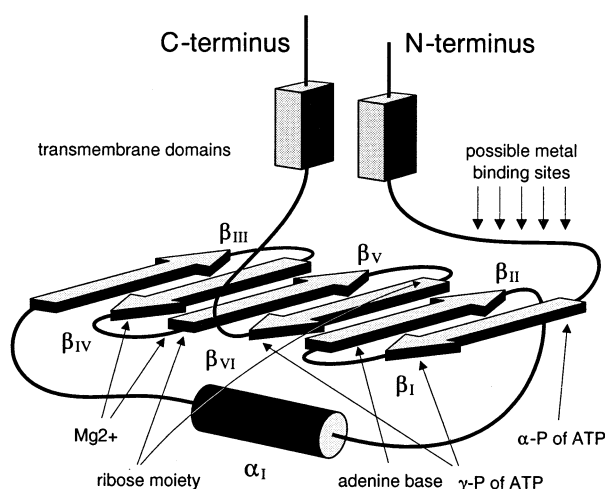


Fig. 2. Proposed structure for P2X channel proteins containing two transmembrane domains, a six-stranded antiparallel β -pleated sheet and metal binding sites. The ATP binding region may be constructed in an analogous fashion as those of class II aminoacyl-tRNA synthetases. Regions responsible for binding of the different parts of the ATP molecule, and Mg^{2+} ion, and eventually for a further metal ion are indicated by arrows.

primordial forms in the earliest stages of biological evolution (compare [34]). If P2X channel proteins are related to class II aminoacyl-tRNA synthetases, high sequence homology cannot be expected. Additionally, those regions of the enzyme molecules which are responsible for recognition of tRNAs and amino acids may have been lost during evolution of channel proteins. There are indeed some indications for this assumption which will be discussed in a later more detailed paper. Although primary structures of proteins imply common features of tertiary structures, the latter ones are far more stable than primary structures in homologous protein series [35]. Therefore more similarities in tertiary and basic secondary structures of class II synthetase and P2X channels should be expected than in sequence homologies.

The sequence stretches of aminoacyl-tRNA synthetases aligned for any β -sheet strand normally exhibit a general common structure together with individual variations, whereas the channel proteins are more conserved. In many cases the sequence stretches are even completely identical. This may be due to a different degree of evolution, and is also typical for sequences important for the functions of the respective protein.

The sequence of the channel proteins, which may form a β -pleated sheet consisting of six strands (positions 170–330, about 160 residues, placed between the cysteine-rich region (110–170) and the second transmembrane domain), has a length sufficient for such a structural feature. The N-terminal cysteine-rich region exhibits the general sequence: $CX_{8-10}CX_5CX_{14-16}CX_{9-10}CX_5C$. Searches carried out with a GCG software package in protein data bases show a considerable high sequence similarity of this sequence region with metallothioneins. As a second possibility some of the conserved cysteine residues can also be regarded as regions of zinc finger motifs in which stretches of amino acids are folded around a zinc ion. The side chains of four amino acids, cysteine or histidine residues, are always coordinated with the zinc ion [36]. Similar zinc binding motifs are also found in both classes of aminoacyl-tRNA synthetases (listed in [37]), and the P2X channel protein may very well possess a metal binding site in the mentioned region (Fig. 2).

For all ATP binding motifs of the synthetases, corresponding sequences can be found in the channel proteins (Figs. 1 and 2): in the β_I region an arginine may bind to the α -phosphate moiety. In the β_I and β_{VI} regions, arginine, lysine, or histidine may interact with the γ -phosphate group. In the β_{II} region a phenylalanine may be responsible for stacking interactions with the adenine base, a glycine in β -sheet strand β_{VI} for interaction with the ribose, and a glutamate at the end of β_{IV} may bind the magnesium cation of an $[ATP-Mg]^{2-}$ complex, together with a second glutamate or aspartate between β_{IV} and β_V (not shown in Fig. 1), which is also involved in interaction with the ribose moiety.

Remarkably, the channel region around position 150, which corresponds to synthetase motif 1, does not show very convincing analogies (not shown). In the enzymes the sequences of motif 1 are responsible for aggregation of two or four protein chains to the holoenzyme. The region around position 330 (termed H5 or P-region) has been suspected to be involved in channel pore formation and is characterised by a sequence motif FX_8GX_5P [38,39]. It exhibits certain similarity to motif 1 of the synthetases which is characterised as $G(FY)X_5P$. The motif 1-like structure may possibly be used

by the channel proteins for aggregation of subunits and formation of a ion-pore structure. For example, in the crystal structure of phenylalanyl-tRNA synthetase from *T. thermophilus* [26] four subunits $(\alpha\beta)_2$ aggregate with their motif 1 regions to a 'quasi barrel'. Possibly, the primordial motif 1 occurred again by gene duplication mechanisms during evolution of P2X channels behind motif 3 and was used for construction of ion pores. In this context it should be mentioned that voltage-gated ion channels have a tetrameric structure, confirmed by the first published crystal structure of a K^+ channel [40].

Although the sequence similarities indicate perhaps a six-stranded β -pleated sheet for the ATP binding domain of the channel proteins, details of their tertiary structures may deviate from that of the synthetase core. However, our hypothesis may be helpful for future research including X-ray analyses of channel protein crystals, NMR spectroscopy and modelling of the P2X receptor ATP binding site, or affinity labelling experiments.

Acknowledgements: The authors would like to thank Dr. Florentina Soto for helpful comments and Dr. Gillian Busch for carefully reading the manuscript.

References

- [1] Brake, A.J. and Julius, D. (1996) *Annu. Rev. Cell Dev. Biol.* 12, 519–541.
- [2] Buell, G., Collo, G. and Rassendren, F. (1996) *Eur. J. Neurosci.* 8, 2221–2228.
- [3] Soto, F., Garcia-Guzman, M. and Stühmer, W. (1997) *J. Membr. Biol.* 160, 91–100.
- [4] Eriani, G., Delarue, M., Poch, O., Gangloff, J. and Moras, D. (1990) *Nature* 347, 203–206.
- [5] Eriani, G., Cavarelli, J., Martin, F., Ador, L., Rees, B., Thierry, J.-C., Gangloff, J. and Moras, D. (1995) *J. Mol. Evol.* 40, 499–508.
- [6] Schimmel, P. (1991) *Trends Biochem. Sci.* 16, 1–3.
- [7] Burbaum, J.J. and Schimmel, P. (1991) *J. Biol. Chem.* 266, 15965–15968.
- [8] Moras, D. (1993) *Biochimie* 75, 651–657.
- [9] Delarue, M. and Moras, D. (1993) *BioEssays* 15, 675–687.
- [10] Leberman, R., Härtlein, M. and Cusack, S. (1991) *Biochim. Biophys. Acta* 1089, 287–298.
- [11] Härtlein, M. and Cusack, S. (1995) *J. Mol. Evol.* 40, 519–530.
- [12] Åberg, A., Yaremchuk, A., Tukalo, M., Rasmussen, B. and Cusack, S. (1997) *Biochemistry* 36, 3084–3094.
- [13] Fujinaga, M., Berthet-Colominas, C., Yaremchuk, A.D., Tukalo, M.A. and Cusack, S. (1993) *J. Mol. Biol.* 234, 222–233.
- [14] Belrhali, H., Yaremchuk, A., Tukalo, M., Larsen, K., Berthet-Colominas, C., Leberman, R., Beijer, B., Sproat, B., Als-Nielsen, J., Grubel, G., Legrand, J.-F., Lehmann, M. and Cusack, S. (1994) *Science* 263, 1432–1436.
- [15] Belrhali, H., Yaremchuk, A., Tukalo, M., Berthet-Colominas, C., Rasmussen, B., Bösecke, P., Diat, O. and Cusack, S. (1995) *Structure* 3, 341–352.
- [16] Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N. and Leberman, R. (1990) *Nature* 347, 249–255.
- [17] Cusack, S., Yaremchuk, A. and Tukalo, M. (1996) *EMBO J.* 15, 2834–2842.
- [18] Arnez, J.G., Harris, D.C., Mitschler, A., Rees, B., Francklyn, C.S. and Moras, D. (1995) *EMBO J.* 14, 4143–4155.
- [19] Delarue, M., Poterszman, A., Nikonov, S., Garber, M., Moras, D. and Thierry, J.C. (1994) *EMBO J.* 13, 3219–3229.
- [20] Poterszman, A., Delarue, M., Thierry, J.-C. and Moras, D. (1994) *J. Mol. Biol.* 244, 158–167.
- [21] Ruff, M., Krishnaswamy, S., Boeglin, M., Poterszman, A., Mitschler, A., Podjarny, A., Rees, B., Thierry, J.C. and Moras, D. (1991) *Science* 252, 1682–1689.

- [22] Cavarelli, J., Rees, B., Ruff, M., Thierry, J.-C. and Moras, D. (1993) *Nature* 362, 181–184.
- [23] Cavarelli, J., Eriani, G., Rees, B., Ruff, M., Boeglin, M., Mitschler, A., Martin, F., Gangloff, J., Thierry, J.-C. and Moras, D. (1994) *EMBO J.* 13, 327–337.
- [24] Logan, D.T., Mazauric, M.-H., Kern, D. and Moras, D. (1995) *EMBO J.* 14, 4156–4167.
- [25] Mosyak, L. and Safro, M. (1993) *Biochimie* 75, 1091–1098.
- [26] Mosyak, L., Reshetnikova, L., Goldgur, Y., Delarue, M. and Safro, M.G. (1995) *Nature Struct. Biol.* 2, 537–547.
- [27] Goldgur, Y., Mosyak, L., Reshetnikova, L., Ankilova, V., Lavrik, O., Khodyreva, S. and Safro, M. (1997) *Structure* 5, 59–68.
- [28] Onesti, S., Miller, A.D. and Brick, P. (1995) *Structure* 3, 163–176.
- [29] Chou, P.Y. and Fasman, G.D. (1978) *Adv. Enzymol.* 47, 45–148.
- [30] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) *J. Mol. Biol.* 120, 97–120.
- [31] Kneller, D.G., Cohen, F.E. and Langridge, R. (1990) *J. Mol. Biol.* 214, 171–182.
- [32] Rost, B. and Sander, C. (1993) *J. Mol. Biol.* 232, 584–599.
- [33] Salamov, A.A. and Solovyev, V.V. (1997) *J. Mol. Biol.* 268, 31–36.
- [34] Eigen, M. (1971) *Naturwissenschaften* 58, 465–523.
- [35] Zuckerkandl, E. (1975) *J. Mol. Evol.* 7, 1–57.
- [36] Rhodes, D. and Klug, A. (1993) *Sci. Am.* 268, 32–39.
- [37] Nureki, O., Kohno, T., Sakamoto, K., Miyazawa, T. and Yokoyama, S. (1993) *J. Biol. Chem.* 268, 15368–15373.
- [38] Brake, A.J., Wagenbach, M.J. and Julius, D. (1994) *Nature* 371, 519–523.
- [39] North, R.A. (1996) *Ciba Found. Symp.* 198, 91–109.
- [40] Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T. and Mackinnon, R. (1998) *Science* 280, 69–77.